

Metaethics and the Overlapping Consensus

I. INTRODUCTION

One of the central problems of liberalism is the enforcement of moral rules on those who disagree. Liberalism adopts the ideal that individuals should be free to express themselves insofar as possible,¹ and this ideal is in conflict with government enforcing a moral code. The dilemma of enforcing ideas on those who do not agree has been called the paradox of liberalism.² Enforcing morality is not just a dilemma for liberalism, however, but a question that any governmental system must address.

If morality is subjective—if moral truths do not have an existence independent of any individual human thought process—then justifying the selection of one moral precept (one person's thought process) over another cannot be done objectively. However, if moral truths have an independent existence, the enforcement of morals can be justified by this objective correctness. For example, if abortion is an objective moral wrong, the person who believes that abortion is morally acceptable is simply mistaken. If there is no objective moral right or wrong, no moral view can be declared by society to be objectively correct, and society will have to enforce moral rules based on the subjective assessments of the controlling members of the society.³

This debate about whether moral systems are objective or subjective is a "metaethical" debate—a debate, not about the substance of morals themselves, but about what we can know concerning morals.⁴ Examples of metaethical questions are whether we can know that any moral ideas are true, how we can

¹ This is not to imply that there is agreement about what liberalism is or should be. The author believes that the core concept of liberalism is autonomy of the person. This idea cuts across the traditional division between liberal and conservative. For example, libertarians, who are on the right-wing end of the political spectrum in many respects, value autonomy highly. Perhaps the main controversy among those who favor autonomy is the relationship between autonomy and equality. For a survey of some contemporary liberal theories and their critics, see WILL KYMLICKA, *CONTEMPORARY POLITICAL PHILOSOPHY* (1990).

² See, e.g., Jean Hampton, *Should Political Philosophy Be Done without Metaphysics?*, 99 *ETHICS* 791 (1989); Thomas Nagel, *Moral Conflict and Political Legitimacy*, 16 *PHIL. & PUB. AFF.* 215, 215 (1987) ("Robert Frost defined a liberal as someone who can't take his own side in an argument.").

³ Thomas Nagel asks the question whether a political result is universally acceptable because it is right or right because it is universally acceptable. Nagel, *supra* note 2, at 219.

⁴ DAVID O. BRINK, *MORAL REALISM AND THE FOUNDATIONS OF ETHICS* 1–2 (1989).

best justify moral beliefs, and how moral ideas cause or motivate actions.⁵ No agreement exists among philosophers on any correct metaethical theory, and in particular no agreement exists on the question of whether morality is objective or subjective.⁶

In the last several years there has been an increasing interest in and debate about metaethics. This debate has affected not only general philosophy but also the philosophy of law. Theories of justice have been articulated in the last twenty years that owe a great deal to the metaethical debate.⁷ Theories of justice can have practical consequences in areas of social policy such as public education, distribution of wealth, church-state relations, and punishment of criminals.⁸ The discussion of theories of justice in this Comment uses several examples related to the question of abortion and mentions other social questions such as the rights of children and animals. These examples show the practical consequences of metaethical theories.

John Rawls, in his theory of justice which he calls "justice as fairness,"⁹ attempts to deal with metaethical questions and the associated liberal paradox by defining a framework for government action that is neutral, insofar as

⁵ See, e.g., H.J. McCLOSKEY, *META-ETHICS AND NORMATIVE ETHICS* 1 (1969) ("We enter the sphere of meta-ethics when we reflect about what we are doing when we make a moral judgment . . .").

⁶ In this Comment, I will assume that realist theories support the idea of objective truth and that antirealist theories support the idea that moral truths cannot be objectively proven. For an argument that objectivity is not an essential component of realism, see BRINK, *supra* note 4, at 125-30. I do not mean by "objective truth" to exclude all types of relativism from moral realism. For a discussion of types of relativist ideas that can be encompassed within a definition of objective moral realism, see ALAN H. GOLDMAN, *MORAL KNOWLEDGE* 8-9 (1990).

⁷ See, e.g., BRUCE A. ACKERMAN, *SOCIAL JUSTICE IN THE LIBERAL STATE* (1980); KENT GREENAWALT, *RELIGIOUS CONVICTIONS AND POLITICAL CHOICE* (1988); JOHN RAWLS, *A THEORY OF JUSTICE* (1971).

⁸ In recent years there has been an increase in interest in the philosophy of law. In particular, there are philosophers who write about law, such as Robert Audi, Richard Brandt, David Brink, William Galston, Jean Hampton, Thomas Nagel, and John Rawls; and lawyers who write about philosophy, such as Bruce Ackerman, Ronald Dworkin, Edward Foley, Kent Greenawalt, Michael Perry, Richard Posner, and Lawrence Solum. The intersection of law and philosophy has thus far perhaps had its greatest influence in law in the area of the relationship between government and religion and specifically in the interpretation of the religion clauses of the First Amendment. This type of analysis, however, is equally applicable to other areas of law, such as those broadly included under the rubric of laws affecting social justice.

⁹ JOHN RAWLS, *POLITICAL LIBERALISM* (1993).

possible,¹⁰ between the various metaethical theories.¹¹ One of the ideas in Rawlsian liberalism is that people who hold different metaethical views can agree on governmental policies, meaning that an area of "overlapping consensus" can exist in spite of metaethical differences.¹² Rawls defines a detailed conception of the person¹³ which he uses to derive an idea of "public reason"¹⁴ made up of common values on which all reasonable members of society can agree.¹⁵ Thus, Rawls uses his concepts of the person and public reason to exclude illiberal ideas, which he defines as unreasonable ideas, from the overlapping consensus.¹⁶ This is Rawls's solution to the liberal paradox.

If government tries to be neutral to all metaethical theories, government cannot make value judgments and if government cannot make value judgments then excluding practices such as slavery and religious intolerance from the overlapping consensus becomes difficult.¹⁷ The difficulty with solving the liberal paradox by Rawls's conception of the person is that, because it is a constructed definition, it is open to the charge of contingency. The difficulties with Rawls's definition of public reason are twofold. First, it is based on a contingent definition of the person, which, since it does not claim to be grounded in truth, cannot be justified objectively. Second, reason is not an adequate basis for selecting public reason since reasonable people can disagree about many values.

This Comment argues that metaethical neutrality is counter to the nature of government and that the conflicts Rawls has taken out at the metaethical level (by means of governmental neutrality to metaethical theories) are put back in at the ethical level with the concepts of the constructive definition of the person and Rawls's definition of public reason. This Comment proposes a modified overlapping consensus that is neutral only between well-justified metaethical theories and that can be used to identify public reasons on a metaethical basis, rather than on a merely contingent basis.

¹⁰ *Id.* at 95, 133-72; John Rawls, *The Idea of an Overlapping Consensus*, 7 OXFORD J. LEGAL STUD. 1 (1987).

¹¹ RAWLS, *supra* note 9, at 47-66; John Rawls, *Justice as Fairness: Political Not Metaphysical*, 14 PHIL. & PUB. AFF. 223 (1985).

¹² RAWLS, *supra* note 9, at 131-72.

¹³ John Rawls, *Kantian Constructivism in Moral Theory*, 77 J. PHIL. 515 (1980).

¹⁴ For an interesting discussion of the development of the term "public reason," see Laurence B. Solum, *Constructing an Ideal of Public Reason*, 30 SAN DIEGO L. REV. (forthcoming 1994).

¹⁵ RAWLS, *supra* note 9, at 223-27.

¹⁶ *Id.* at 36-37, 226-27, 246-47.

¹⁷ WILLIAM A. GALSTON, LIBERAL PURPOSES: GOODS, VIRTUES, AND DIVERSITY IN THE LIBERAL STATE 79-117 (1991); Hampton, *supra* note 2.

Because governmental action and personal action are different, it is not desirable for government to be completely neutral between metaethical theories. The nature of the difference between governmental and personal action consists in the fact that government orders relations between people and therefore should function on the basis of widely accessible concepts. Individuals may choose to evaluate metaethical theories on the basis of truth alone, but a government must evaluate metaethical theories on the basis of both truth and justification.¹⁸ A metaethical theory can have as good a claim to truth as any other metaethical theory, while not having as good a claim to being justified. Well-justified metaethical theories are those that have a wide circle of justification. A new type of overlapping consensus that excludes only metaethical theories that are inadequately justified results in the exclusion of so-called "illiberal" ideas without resorting to an artificially constructed definition of the person. This Comment concludes that government can be neutral between realism and antirealism but should not be neutral between well-justified and poorly justified metaethical theories.¹⁹

Part II describes the major metaethical theories, gives an overview of their development and the arguments for and against their truth, and concludes that the overlapping consensus should exclude empirical foundationalism. Part III examines and evaluates the justification of various metaethical theories and determines that the overlapping consensus should exclude intuitionism. Part IV describes and criticizes Rawls's idea of the overlapping consensus and proposes a new type of overlapping consensus. This new overlapping consensus is constructed to be neutral only between the realist metaethical theory of coherentism and the antirealist metaethical theory of constructivism.

II. METAETHICAL THEORIES AND TRUTH

A. *Realism and Antirealism*

What is realism? Realism has to do with the belief that there are truths that exist "out there." That is, realists believe that there are real things or ideas that exist in the world. An example of a realist theory is the jurisprudential theory

¹⁸ The difference between truth and justification is discussed in Part III.

¹⁹ This paper is written for the nonphilosopher, and thus there is considerable oversimplification of some of the most complicated philosophical questions. However, an attempt has been made to simplify matters without doing violence to the core concepts. Interested readers are advised to consult the cited works for more rigorous and detailed explications of the various philosophical theories.

of natural law.²⁰ This theory holds that the law has its own existence separate from humanity—humans do not invent the law; they just discover or describe it. The implication of this theory is that there are right answers to every question. For example, a believer in natural law would say that there is one right answer to the question of whether abortions should be legal or not and that right answer exists in spite of what humans do or think.

A more formal way of defining realism is to say that realism is the belief that there are things that exist independently of any person's mind or any person's beliefs or attitudes.²¹ A problem with this definition, however, is that it would mean our thoughts and psychological states are not real.²² Rewording this definition, we can arrive at a minimal definition of realism such as: "there are facts of a certain kind which are independent of our evidence for them."²³

Realism is contrasted with antirealism, which in its most extreme skeptical forms says that there are no true facts.²⁴ Milder forms of skepticism say that there may be true facts, but there are no facts that we can know to be true or know to be true with certainty.²⁵

It is possible to be a realist in terms of observational facts about the world in general while being an antirealist about the existence of objective moral truths.²⁶ That is, it is possible to believe that the things we perceive or the proofs of science are true and describe a reality that exists independently of our evidence for them and at the same time believe that there is no one "true" answer to moral questions such as whether abortion is wrong.²⁷ The difference

²⁰ See generally Suzanna Sherry, *The Founder's Unwritten Constitution*, 54 U. CHI. L. REV. 1127 (1987).

²¹ GOLDMAN, *supra* note 6, at 7.

²² BRINK, *supra* note 4, at 15.

²³ *Id.*

²⁴ 2 RENE DESCARTES, *Meditations on First Philosophy*, in THE PHILOSOPHICAL WRITINGS OF DESCARTES 12 (John Cottingham et al. trans., 1984) [hereinafter DESCARTES, *Meditations*]; 1 RENE DESCARTES, *Discourse on the Method*, THE PHILOSOPHICAL WRITINGS OF DESCARTES 111 (John Cottingham et al. trans., 1984) [hereinafter DESCARTES, *Discourse*].

²⁵ See generally George S. Pappas & Marshall Swain, *Introduction*, in ESSAYS ON KNOWLEDGE AND JUSTIFICATION 11, 37 (George S. Pappas & Marshall Swain eds., 1978).

²⁶ BRINK, *supra* note 4, at 5–7.

²⁷ Even if a person believes that there are no moral truths which are comparable to observational truths, that person may believe that government can penalize certain types of actions on moral grounds. A moral antirealist may believe that there is a right answer and a wrong answer to the question of whether abortion should be legal, but the moral antirealist will believe that this answer is right or wrong because of human attitudes and beliefs, not because of an independently existing moral structure. See J.L. MACKIE, *ETHICS: INVENTING RIGHT AND WRONG* 25–27 (Penguin Books 1990) (1977).

between metaethics and metaphysics is that metaethics concerns morals while metaphysics concerns the world in general. Although this Comment is concerned with metaethics, general realist and antirealist arguments will also be discussed because metaethical and metaphysical arguments are closely related.

Realism today can be divided into two broad types, which are known as foundationalism and coherentism. Antirealism today can be divided into two broad types known as skepticism and constructivism. Foundationalism was the original version of realism, and skepticism was the original version of antirealism. One way of thinking about the metaethical debate is to view the battle between foundationalism and skepticism as giving rise to the more recent theories of coherentism and constructivism. Coherentism is the realist's way of meeting the skeptical challenge, and constructivism is the antirealist's way of meeting the coherentist defense of realism. Part II-A examines the traditional theories of foundationalism, intuitionism, and skepticism in sub-parts one through four. Part II-B will examine the theories of coherentism and constructivism.

1. *Foundational Realism*

Foundational realists claim that true facts exist and that we can determine these true facts by tracing our conclusions and beliefs back to foundational truths.²⁸ There are many different types of foundational realism,²⁹ including an extreme version which holds that foundationalism need not guarantee any type of objective truth.³⁰ A detailed description of all of the types of

²⁸ See NICHOLAS RESCHER, *THE COHERENCE THEORY OF TRUTH* 317 (1982) ("Within the epistemic structure of our knowledge of truth, such basic truths are to serve as a foundation; other truths are made to rest upon them, but they rest on no others: like the axioms of a deductive system they provide the ultimate support for the entire structure.").

²⁹ See, e.g., ARISTOTLE, *Metaphysica*, in *BASIC WORKS OF ARISTOTLE* 261 (Richard McKeon ed., 1941); RODERICK M. CHISHOLM, *THE FOUNDATIONS OF KNOWING* (1982); James W. Cornman, *Foundational versus Nonfoundational Theories of Empirical Justification*, in *ESSAYS ON KNOWLEDGE AND JUSTIFICATION*, *supra* note 25, at 229; Mark Pastin, *Modest Foundationalism and Self-Warrant*, in *ESSAYS ON KNOWLEDGE AND JUSTIFICATION*, *supra* note 25, at 279 (Pastin makes an important distinction between foundationalism, which holds that foundational beliefs must be true in all situations at all times, and between "radical foundationalism" and "modest foundationalism," which hold that a foundational belief need only be self-justifying for a particular person at a particular time); 1 *THE PHILOSOPHICAL WRITINGS OF DESCARTES*, *supra* note 24; 2 *THE PHILOSOPHICAL WRITINGS OF DESCARTES*, *supra* note 24.

³⁰ ALVIN I. GOLDMAN, *EPISTEMOLOGY AND COGNITION* 194-98 (1986).

foundationalism is beyond the scope of this Comment.³¹ However, most, if not all, types of foundationalism contain the core idea that any particular belief that a person is justified in holding must be traceable to one or more "foundational" beliefs that are capable of justifying the particular belief.³² The foundational beliefs themselves are a special class of beliefs that are self-justifying. "Strong" foundationalism holds further that foundational beliefs are evidence for truth, not just evidence for justification in believing the truth.³³

An important traditional version of foundationalism held that beliefs derived from *experience* were self-justifying.³⁴ That is, if one perceived that a book was on the table, that perception would be a justification for the belief. Experiential foundationalists believed observations could be taken as true. However, as time went on, many philosophers became convinced that perceptions can be unreliable and that science itself must assume certain truths without the ability to prove them, and thus science and perceptions are ultimately based on assumptions and not on foundational truths.³⁵ One of the important objections that led to the view that observations cannot constitute foundational truths was the skeptical argument of the infinite regress.

³¹ The reader versed in philosophy will also note that in this discussion I elide over the difference (if any) between deductive and inductive inferences and the related distinction (if any) between a priori and a posteriori knowledge. For discussions of representative positions on these questions, see LAURENCE BONJOUR, *THE STRUCTURE OF EMPIRICAL KNOWLEDGE* 191–211 (1985); JONATHAN DANCY, *AN INTRODUCTION TO CONTEMPORARY EPISTEMOLOGY* 212–26 (1985); ALAN H. GOLDMAN, *EMPIRICAL KNOWLEDGE* 189–215 (1991); W.V. QUINE, *PURSUIT OF TRUTH* 1–36 (1992); RICHARD RORTY, *PHILOSOPHY AND THE MIRROR OF NATURE* (1979).

³² GOLDMAN, *supra* note 30, at 79.

³³ Stronger foundationalisms would endorse all of the following theses: (1) there are immediately justified (validated) beliefs, beliefs justified without appeal to evidence; (2) the validation of all other beliefs must terminate in one or more immediately validated beliefs (beliefs cannot be supported solely by other beliefs, unless the latter are foundational); and (3) foundational beliefs are infallible, incorrigible, or self-evident. GOLDMAN, *supra* note 31, at 136.

³⁴ For a discussion of perceptual beliefs, see *id.* at 71–80.

³⁵ See, e.g., ALFRED J. AYER, *LANGUAGE TRUTH AND LOGIC* 90–95 (1946); LUDWIG WITTGENSTEIN, *CULTURE AND VALUE* (Peter Winch trans., 1984).

2. *The Skeptical Argument of the Infinite Regress*

The skeptical arguments generally fall into two broad categories.³⁶ One is the argument from error, which will be addressed below.³⁷ The second is the argument that all truths must be justified by other truths, leading to an infinite regress that can never terminate in foundational truths.³⁸

A good way of posing this problem is to imitate a child who asks “why?” after every answer. Every reason we have for a belief can be questioned and must be justified by another belief which must have its own reason and so on. This process will go on infinitely, and we can therefore never justify any belief unless we can find some type of foundational beliefs that either do not need to be justified or are self-justifying.

Realists believe that there is a difference between “truth” and “justification.”³⁹ Justification, roughly speaking, is the factual standard for believing something to be true,⁴⁰ meaning what reason or reasons we have for believing the truth of a particular idea. For example, the belief that the earth is round can be justified by satellite photographs. The problem with this type of justification is that it can go on forever. If we rely on the satellite photographs as justification that the earth is round, we must justify our belief that the satellite photographs are accurate and so forth. So, although we may have a justification for believing something, this does not mean that it is true. Whatever reason I provide for believing something, I can always ask “why believe that reason?” and demand yet another justification. This is why the problem is called the *infinite regress* problem.

Responses to the idea of the infinite regress generally fall into one of two categories. First, the justificatory chain can be allowed to loop back on itself in some fashion, thus ultimately justifying truths by other truths. This idea is discussed below in connection with coherentism and constructivism.⁴¹ The second idea is that a justificatory chain can be terminated in a self-evident belief. This is the idea of intuitionism.

³⁶ GOLDMAN, *supra* note 30, at 28–41. Goldman identifies three grounds for skepticism, which I collapse into two. Goldman’s three grounds are (1) error, (2) infinite regress, and (3) rival-hypothesis. Goldman suggests that the rival-hypothesis problem and the error problem can be dealt with together, and I agree.

³⁷ See *infra* notes 58–73 and accompanying text.

³⁸ BONJOUR, *supra* note 31, at 17–18.

³⁹ BRINK, *supra* note 4, at 31–36.

⁴⁰ GOLDMAN, *supra* note 30, at 23–24.

⁴¹ See *infra* notes 74–83 and accompanying text.

3. Intuitionism

Beginning at the turn of the century, realists elaborated the idea of intuitionism as a response to the skeptical criticism of the infinite regress.⁴² Intuitionism is a type of foundationalism, in that intuitionism accepts the idea that there are self-justifying beliefs. However, intuitionism seeks to avoid the problem of the infinite regress by asserting that humans have a special faculty for determining the truth of foundational beliefs. This special faculty is intuition.⁴³ Truths derived by intuition are present to us in an immediate way that bypasses the need for any justification.⁴⁴ Intuitionists have been among the main defenders of moral realism⁴⁵ although intuitionism can also be conceptualized as a theory of general realism.

There are several varieties of intuitionism, but the classic types were elaborated by Henry Sidgwick.⁴⁶ Sidgwick divided intuitionist theories into three categories. The simplest is "sense intuitionism"⁴⁷ or "perceptual intuitionism," which relies on the individual's feelings as determinants of truth.⁴⁸ Sidgwick describes this form as "recogni[zing] simple immediate intuitions alone and discard[ing] as superfluous all modes of reasoning to moral conclusions."⁴⁹ Sidgwick's second category of intuitionism is "dogmatic intuitionism." Dogmatic intuitionism holds that we ordinarily have a fairly good intuitive sense of morals but that to state moral truths precisely "requires a special habit of contemplating clearly and steadily abstract moral notions."⁵⁰ Sidgwick calls this the "morality of common sense."⁵¹

⁴² GEORGE E. MOORE, *PRINCIPIA ETHICA* (1922); W.D. ROSS, *THE RIGHT AND THE GOOD* (Hackett Pub. Co. 1988) (1930); HENRY SIDGWICK, *THE METHODS OF ETHICS* (Hackett Pub. Co. 1981) (1907).

⁴³ See, e.g., ROSS, *supra* note 42, at 75–133.

⁴⁴ RESCHER, *supra* note 28, at 207 ("There must be a starter set of *primitive* (ungrounded, immediate, 'intuitive') truths and, outside this special category, truths can only be established from or grounded upon other truths.").

⁴⁵ BRINK, *supra* note 4, at 2–3.

⁴⁶ SIDGWICK, *supra* note 42, at 96–104.

⁴⁷ Francis Hutcheson, *An Essay on the Nature and Conduct of the Passions and Affections*, in 1 *BRITISH MORALISTS 1650–1800*, at 303, 303–73 (D.D. Raphael, ed., Hackett Pub. Co. 1991) (1969).

⁴⁸ SIDGWICK, *supra* note 42, at 97.

⁴⁹ *Id.* at 100.

⁵⁰ *Id.* at 101.

⁵¹ *Id.* at 101–02.

Sidgwick criticizes both perceptual intuitionism and dogmatic intuitionism on the grounds that it is not evident why we should accept them as true.⁵² Sidgwick attempts to answer this criticism by defining a third category of intuitionism, "philosophical intuitionism." Philosophical intuitionism, "while accepting the morality of common sense as in the main sound, still attempts to find for it a philosophical basis which it does not itself offer."⁵³

Intuitionism, especially moral intuitionism, has been criticized (as Sidgwick recognized) on the basis that there is no agreement among humans on these allegedly self-evident foundational beliefs.⁵⁴ A position on the rightness or wrongness of abortion, for example, does not seem to be self-evident. An additional, related criticism is that intuitionism assumes the existence of moral experts⁵⁵ because the fact that no general agreement on moral truths exists among humans means that humans must have varying degrees of moral intuition.⁵⁶

The idea of intuitive truths can be seen as an example of the worst sort of circularity—a concept justifying itself by itself.⁵⁷ But what about Sidgwick's third type of intuitionism—philosophical intuitionism? It is debatable to what extent most scholars today would consider this to be true intuitionism. The further an intuitionist moves along the scale of identifying or justifying beliefs by reason, the further the intuitionist moves away from the idea of self-justifying truths.

Because intuitionism posits a special faculty of human knowing, intuitionism does successfully meet the argument of the infinite regress. One may or may not agree that humans have this special faculty, but if this special faculty does exist it would solve the problem of the infinite regress. In addition to the skeptical problem of the infinite regress, however, intuitionism must also confront the skeptical argument from error.

4. *The Skeptical Argument From Error*

One skeptical idea is that we can never know anything because we can never be free from the possibility of error.⁵⁸ The most extreme form of this

⁵² *Id.* at 100–02.

⁵³ *Id.* at 102.

⁵⁴ See, e.g., RICHARD B. BRANDT, *ETHICAL THEORY* 192–96 (1959); MACKIE, *supra* note 27, at 36.

⁵⁵ BRINK, *supra* note 4, at 95–98.

⁵⁶ See, e.g., J.B. Schneewind, *Natural Law, Skepticism, and Methods of Ethics*, 52 J. HIST. IDEAS 289 (1991).

⁵⁷ BRINK, *supra* note 4, at 116–17.

⁵⁸ ANTHONY O'HEAR, *WHAT PHILOSOPHY IS* 111–19 (1985).

argument says that a person cannot even know that anything outside of her own mind exists.⁵⁹ For example, we could be under the control of demons,⁶⁰ or we could be nothing more than a brain in a vat, hooked up to wires that feed us illusions of all that we think we perceive.⁶¹

One response to the brain-in-the-vat problem is to ask whether it is a real problem. That is, if I am nothing but a brain in a vat, and an illusion that there is an outside world has been created that is so perfect that I can never detect its falsity, perhaps there is no difference for me between the truth and the falsity of the system, and thus the problem loses its meaning.⁶²

Another way to look at this is by analogy to the "many universes" hypothesis in physics.⁶³ One proposed explanation for some of the mysteries of quantum physics is that there are actually an infinite number of parallel universes where every possible physical outcome is played out. Possibly, if these universes exist, they do not have any contact points with our universe—these other universes might be totally outside of our ken, and there might be no way we could ever know them. If this is the case, it is reasonable to ask if these universes really exist for us. Certainly they do not exist in the way that we think of ordinary things as existing. It might not be quite accurate to call other universes "false," but if we can never have any contact with them, they do not seem to fit our concept of "true" since we can never verify any qualities they may have. Things that are totally outside of our ken might be neither true nor false—they might be simply meaningless to us.

One could say that whether or not the brain-in-the-vat problem is a meaningful problem will depend on whether any situations can be identified where the results of a person's experience would be different if the person were a brain in a vat or a human in a world of other humans. Remember that one of the starting conditions of the brain-in-the-vat problem is that the illusion is so perfect that we can never detect it. If we could detect it, we would know the truth and that particular skeptical possibility would be defeated.

In ordinary situations, whether or not a person is a brain in a vat makes no difference. We will still eat to avoid hunger, whether hunger is an illusion or not. But one could argue that a person who is a brain in a vat would have a reason to be completely selfish. That is, if a brain in a vat is the only thing that

⁵⁹ GEORGE BERKELEY, *A TREATISE CONCERNING THE PRINCIPLES OF HUMAN KNOWLEDGE* 45–146 (G.J. Warnock ed., Open Court 1986) (1710).

⁶⁰ DESCARTES, *Meditations*, *supra* note 24, at 15.

⁶¹ HILARY PUTNAM, *REASON, TRUTH AND HISTORY* (1981).

⁶² O'HEAR, *supra* note 58, at 111–17.

⁶³ See PETER COVENEY & ROGER HIGHFIELD, *THE ARROW OF TIME* 132–34 (1990) (criticizing the theory); JOHN GRIBBIN, *IN SEARCH OF SCHRODINGER'S CAT* 235–54 (1984) (discussing this theory in nonmathematical terms).

exists, there is no reason for any kind of altruism or consideration for other creatures since they are mere illusions.⁶⁴ However, when you look at actual behaviors, it is hard to detect when there would be any actual difference. For example, even if I am a brain in a vat, I am not likely to steal the belongings of other people for fear of the social ostracism or criminal sanctions that I would suffer, illusion though they may be. Furthermore, I am likely to be considerate of those close to me because I enjoy their company, whether it is an illusion or not. If I am a brain in a vat I may choose not to give money to charities for strangers whom I will never see, but I might decide to do so on the theory that, since the illusion is perfect, the welfare of these others may influence my illusion in some way I cannot predict. In addition, humans who believe they are in the world of humans are not necessarily noted for their altruism to people whom they will never see.

One objection to the argument that a person who is a brain in a vat would be completely selfish is that it presupposes that people who believe they are people in a world with other people have a reason, just because of that fact, to be altruistic. It is not at all obvious why the mere fact that we believe that other people have minds like ours means that we should not be selfish. After all, even if we believe that other people have minds, we still live only in our own mind, not in the mind of another person. The reason for altruism must come from some other source than our belief in the existence of other minds like ours since such a belief is compatible with complete selfishness.

What about risk-taking behavior? If people believe they are a brain in a vat and that their physical body is merely an illusion, will those people be more likely to take physical risks? For example, I might choose to smoke cigarettes in spite of the risk of lung cancer because lung cancer is only an illusion.

In the first place, a perfect illusion will include pain, and this is one reason to avoid physical risks. Secondly, if the illusion is perfect, the illusion must include death as well. Of course, death could be merely an illusion, and if a person who is really a brain in a vat throws their illusory body off a cliff, perhaps they will immediately be transported into a fresh illusion. But if the illusion is perfect, something will surely happen. We will be unsure what, but this is exactly the situation of humans who believe they are humans in the world. Humans in the world do not know exactly what happens to them when they die. It is possible that our minds merely switch to another body. However, humans in the world generally do not want to take that risk, and people who believe they are a brain in a vat would probably not want to either. Perhaps the forces that control the illusions directed to the brain in the vat would continue

⁶⁴ DANCY, *supra* note 31, at 66-82.

the illusion into an illusion of death. It is unlikely that one who suspects he may be a brain in a vat would elect to find out.

A more general objection to the argument that one who believes that he is a brain in a vat might behave differently than one who believes that he is a creature in the world is the definitional consideration that one can never have any evidence that one is a brain in a vat.⁶⁵ Therefore, one is never certain whether one is or is not a creature in the real world. A creature that is uncertain about the nature of reality is not likely to adopt a course of action that would be inconsistent with one of two possible alternatives when there are courses of action that are compatible with both.

These arguments show that it is possible to conclude that the brain-in-the-vat problem does not state a real problem because the result in any situation will be the same in both systems.⁶⁶ However, this argument, which simply defines away this skeptical possibility, seems somehow unsatisfactory. It seems to matter to us whether we are a brain in a vat or not.⁶⁷ Therefore, although the definitional arguments against the skeptical argument from error are very strong, they are not totally satisfactory.⁶⁸

Bertrand Russell and Alfred North Whitehead addressed a similar problem in mathematics in connection with the problem of paradoxes.⁶⁹ In brief, the problem they looked at was whether it is possible to completely define a system. That is, is it possible from *within* any system (a system being a defined set of things that have some relation to each other) to define the parameters of that system in a completely consistent way? This is much the same problem as the brain-in-the-vat problem. The answer that Russell and Whitehead came to was that it is only possible to completely define a system by separating different categories to avoid self-reference. An example of self-reference is the statement "I always lie." I can separate the categories of self and reference to self by talking about someone else. For example, I can say "Paul always lies" and that statement about Paul is either true or not true. However, the statement "I always lie" necessarily can be neither true nor untrue.

Kurt Godel proved this concept mathematically and showed that a mathematical system could only be consistent (non-paradoxical) when viewed

⁶⁵ O'HEAR, *supra* note 58, at 116–17.

⁶⁶ See GOLDMAN, *supra* note 31, at 331–33 (describing this argument).

⁶⁷ BONJOUR, *supra* note 31, at 179–81.

⁶⁸ ROBERT NOZICK, *PHILOSOPHICAL EXPLANATIONS* 197 (1981) ("To think the skeptic overlooks something obvious, to attribute to him a simple mistake or confusion or fallacy, is to refuse to acknowledge the power of his position and the grip it can have on us.").

⁶⁹ ALFRED NORTH WHITEHEAD & BERTRAND RUSSELL, *PRINCIPIA MATHEMATICA* (2d ed. 1927).

from "outside" the system.⁷⁰ That is, a definition of a system must be made independently of the system itself to avoid the paradoxes of self-reference.⁷¹ Godel proved that a mathematical system is inherently incomplete in that there are assertions that can never be either proved or disproved within the system of known mathematics.

The relationship of this mathematical problem to the brain-in-the-vat problem is obvious. Under the terms of the problem as defined, the question of whether we are a brain in a vat can only be answered from outside of the system, since any evidence of an illusion is defined as part of the illusion—as part of the system. Therefore, the question is unanswerable. The truth or falseness of this question seems somehow different from the truth or falseness of ordinary questions that can be proved within the system. The question of the brain-in-the-vat problem seems to be a type of "meta-question" and the answer to it is perhaps true or false in a different way than our ordinary usage of "true" and "false."⁷²

One might object that the brain-in-the-vat problem is not very interesting. This gets back to the question of whether people should care if they are a brain in a vat or not. The important aspect of the brain-in-the-vat problem is not the question itself, however, but rather the fact that it illustrates that there are things that we cannot ever know to be true or false, at least in the ordinary sense of "true" or "false." There are many self-referencing paradoxes scattered through life, like the sentence "I always lie."⁷³ The existence of these paradoxes shows that no metaphysical or metaethical theory can ever establish itself as "true" *except within the terms of a defined system*. This observation is important in the discussion of coherentism and constructivism that follows.

But before looking at coherentism and constructivism, it is helpful to return briefly to the question of how intuitionism survives the skeptical argument from error. As shown in the discussion of the brain-in-the-vat problem, the skeptical argument from error does have some force against intuitionism. However, it seems that intuitionism could still determine truth within the limits of a defined system. For example, if an intuitionist believed in a deistic god, defined that god as the force that started the universe, and made other necessary definitions such as the linearity of time, the skeptical argument from error could only attack the deistic belief by asserting that the system is invalid and could not

⁷⁰ KURT GODEL, ON FORMALLY UNDECIDABLE PROPOSITIONS OF PRINCIPIA MATHEMATICA AND RELATED SYSTEMS (B. Meltzer trans., Dover 1992) (1931).

⁷¹ See also W.V. QUINE, THE WAYS OF PARADOX AND OTHER ESSAYS (1966).

⁷² I believe that Thomas Nagel makes a somewhat similar point regarding morals, if I understand him correctly. See THOMAS NAGEL, THE VIEW FROM NOWHERE (1986).

⁷³ See, e.g., PARADOXES OF RATIONALITY AND COOPERATION (Richmond Campbell & Lanning Snowden eds., 1985); R.M. SAINSBURY, PARADOXES (1988).

attack it as invalid within its defined system. Therefore, the skeptical argument from error does not necessarily destroy the intuitionist claim for at least a certain type of truth.

B. Coherentism and Constructivism

In an attempt to resolve the problem of realism versus antirealism, a number of theorists have devised ideas that fit on a spectrum between the two poles of realism and antirealism. These theories can in general be applied to both the realm of natural sciences and morals. Because the skeptical arguments apply against the natural sciences as well as against ethical inquiry, all realists, not just moral realists, have been involved in this debate.

The in-between positions that have developed fall into two groups. One is called "coherentism," and the other is called "constructivism." Coherentism and constructivism attempt to deal with the skeptical argument from error by defining a sphere within which things can be said to be true. There are many varieties of coherentism and constructivism, and there are also scholars who do not see any crucial difference between coherentism and constructivism. For the purposes of this paper, working definitions of coherentism and constructivism will be given, and the interested reader will be referred to other sources for more detailed discussions.⁷⁴

Briefly, coherentists say that a belief can be justified by the degree to which it fits in with other beliefs to make a coherent whole.⁷⁵ For example, a belief that the earth is round coheres with what one knows of geography, what one hears from other people, and what one has learned from one's own observation that ships on the horizon gradually sink out of sight. Hence, a belief that the earth was flat would conflict with many other beliefs. The explanatory power of a belief is also important in determining how well it coheres with other beliefs. One's belief that the earth is round explains why ships don't fall off the edge of the earth.

Both coherentism and constructivism claim that beliefs are validated by how well they fit into a whole. The difference between realist coherentism and antirealist constructivism is that, according to coherentists, coherence of a belief with other beliefs is considered to be evidence that the belief is

⁷⁴ SIMON BLACKBURN, *SPREADING THE WORD* (1984); RICHARD B. BRANDT, *MORALITY, UTILITARIANISM, AND RIGHTS* (1992); GOLDMAN, *supra* note 6; GILBERT HARMAN, *THE NATURE OF MORALITY* (1977); MACKIE, *supra* note 27; David Copp, *Explanation and Justification in Ethics*, 100 *ETHICS* 237 (1990) (discussing confirmation theory).

⁷⁵ BRINK, *supra* note 4, at 103-04.

objectively true.⁷⁶ Other scholars claim that coherence cannot provide any evidence for objective truth because coherentism is essentially a “house of cards.”⁷⁷ Justifying beliefs by their coherence with other beliefs merely says they are coherent; it doesn’t say they are objectively true. Coherence is not a realist philosophy, these thinkers say, because it does not supply any “ties to the world.”⁷⁸ This objection may be best understood by imagining a complete fantasy world. This world could take place on a different planet at a different time. It could have completely different creatures from any that exist on earth. It could have different physical laws. Such a fantasy could be extremely detailed and complete and fit together perfectly and consistently; yet it will still be completely false because it is just a fantasy. In the same way, a coherentist theory might be completely coherent and still be completely false.

A reply to the objection that a coherent theory may be wildly false is that the coherence of a theory is in itself evidence for objective truth.⁷⁹ This idea can best be explained as a probability analysis which says that the probability that a system is wildly in error decreases as the number of facts that fit into the coherent system increases. Of course, the “truth” of even a very large coherent system is still a relative truth (as explained in the discussion of the brain-in-the-vat problem), but when the size of a relative system approaches the size of the known system, it might be reasonable to claim that relative truth closely approximates the absolute truth of the system.

Moral constructivist thinkers believe that coherence of beliefs is important but that moral beliefs can never be objectively true; they can only have the validity conferred on them by human beliefs and attitudes. In other words, we construct our morality—it is not “out there” anywhere. Moral constructivists believe that a moral system is just a house of cards—merely a human invention. Most moral constructivists do not believe that we are justified in creating just any morality, however, or that a majority is justified in imposing its moral beliefs on a minority merely because they are majority beliefs.⁸⁰ Constructivists generally agree with coherentists that it is important that moral beliefs fit into a coherent whole or conform to a coherent structural system.

Constructivism tries to deal with the skeptical argument from error by maintaining the position that ties to the world are not necessary if you have a

⁷⁶ *Id.* at 125–43. It is important to note, however, that many scholars use the term “coherent” to refer to a system for which they claim no objective ties to truth. To simplify the classification scheme, I am including such views under the “constructivist” rubric.

⁷⁷ See generally PUTNAM, *supra* note 61.

⁷⁸ See, e.g., RICHARD B. BRANDT, A THEORY OF THE GOOD AND THE RIGHT 18–23 (1979).

⁷⁹ See *infra* notes 100–05 and accompanying discussion.

⁸⁰ HARMAN, *supra* note 74, at 65–90.

constructed system. Constructivists concede the objection that the coherence of a system does not provide evidence of truth, and they agree that coherence merely provides evidence of justification. However, most constructivists maintain that what we actually mean by "truth" is "justification."

One could question whether there are any necessary differences between coherentism and constructivism in practice, but there are some important differences in theory. For example, many constructivists believe that the difference between the idea of morals and the idea of experiential things is to some degree semantic.⁸¹ Also, an interesting philosophical problem that was an important factor in the development of constructivism is the problem of how moral facts could drive actions.⁸² Moral realists think of morals as real facts that exist in the world. But people also tend to think of morals as obligations. If morals are just facts in the world, like the fact that the sun rose this morning, then some theorists claim that morals could not impel humans to action. These theorists question how you could have a cause-and-effect link between a fact and an action impelled by obligation—how do facts by themselves produce the impetus to act? And yet a further problem is *why* facts should drive actions. In other words, just because something is, doesn't mean it ought to be. Constructivists say that morals are not facts but are instead the obligations themselves, and thus drive actions directly.

The problems of the action-driving nature of morals and the connection between "is" and "ought" need not be addressed in this paper. Adherents of both coherentist and constructivist theories have put forth reasonable theories about these questions.⁸³ No theory is completely satisfactory, but the answers coherentist and constructivist theorists have put forward are as convincing as the answers to these problems offered by other metaethical theories.

Coherentism and constructivism successfully deal with the skeptical argument of the infinite regress, but do they successfully deal with the skeptical argument from error? As discussed above in connection with the skeptical argument from error, the coherentist and constructivist definition of a system is a strong approach to the skeptical argument from error (and the related subject of the existence of paradoxes), but the definition of a system does not completely answer the skeptical argument from error.

⁸¹ BLACKBURN, *supra* note 74, at 189–96.

⁸² See, e.g., *id.* at 187–89; MACKIE, *supra* note 27, at 27–30.

⁸³ For different views on this question, see BRINK, *supra* note 4, at 81–83 and R.M. HARE, 'Nothing Matters,' in APPLICATIONS OF MORAL PHILOSOPHY 32 (W.D. Hudson ed., 1972).

C. Summary of Metaethical Theories

Thus far, five types of metaethical theories have been described: empirical foundationalism, intuitionism, coherentism, constructivism, and skepticism. Each type is subject to criticisms of its truth, but only empirical foundationalism is conceptually weak in terms of truth. Empirical foundationalism falls to the infinite regress argument which maintains that no beliefs are in fact foundational.

Because I am concerned with *government's* attitude toward metaethical theories, I can also eliminate skepticism as a metaethical theory to be included in the overlapping consensus. Skepticism is not conceptually inferior in terms of truth, but skepticism is a wholly passive approach.⁸⁴ A government could not adopt skepticism as a metaethical policy because government orders the relations among humans; therefore, government is by nature active, not passive. Skepticism cannot provide any structure for such ordering, and thus a government that adopted skepticism would be defining itself out of existence.

One could argue that government should not exist at all and that there is no justification for a belief that human or sentient life is worth preserving or ordering by government. One could say that the fact that skepticism could lead to passivity is an indictment of government, not of skepticism. I agree with the skeptical contention that the idea that government should exist at all requires a leap of faith to a belief in the value of human or sentient life. I cannot answer the skeptical criticism of this leap of faith in a completely satisfactory way. However, the criticism which says that the entire concept of government is arbitrary is a criticism of all possible types of social action, and arguably even inaction can constitute a leap of faith. Even the nihilist who believes that government should not exist because there is no definitive answer to the skeptical argument from error is adopting a position that will have consequences and is therefore adopting a position on social outcomes. It is difficult to understand how we can completely refrain from having a

⁸⁴ See, e.g., O'HEAR, *supra* note 58, at 115-16.

[A] radical scepticism [sic] is a poor policy to adopt, at least if one is interested in any sort of description or categorization of what one experiences. Like the ultimate sceptical doctrine, that there is no truth, scepticism concerning these assumptions is going to leave the sceptic with precious little to say, and if, like the sceptics of Roman times, someone were to regard silence as epistemologically golden, we could point out that we are interested in action as well as in words, and that we need classificatory theories in order to choose future actions and to judge their reasonableness.

government of some type, whether it be by omission or commission. Government is an activity that happens when humans congregate, and it cannot be stopped unless humans agree not to congregate. For this reason, although I cannot answer the skeptical argument from error in a completely satisfactory way, I can set it aside in an evaluation of various types of governmental schemes against each other.

We are therefore left with the metaethical theories of intuitionism, coherentism, and constructivism as theories that all meet the skeptical argument of the infinite regress and are all equally successful (or unsuccessful) at meeting the skeptical argument from error. Part III looks at which of these three metaethical theories survives an analysis based on *justification*.

III. METAETHICAL THEORIES AND JUSTIFICATION

This Part examines the justification of the metaethical schemes whose claims to truth were discussed in Part II. The distinction between truth and justification was referred to briefly in Part II. To review, objective truth is what separates realism and antirealism. Realists believe that there are true facts and that we can know at least some of these true facts.⁸⁵ Antirealists believe either that there are no objectively true facts or that, if true facts exist, we can never know them.⁸⁶ Justification is the process by which we arrive at truth or attempt to arrive at truth.⁸⁷

Moral realists believe that truth and justification are two different things. Moral realists believe that truth exists whether we have any justification for it or not. Justification may provide evidence for truth,⁸⁸ but justification and truth are not identical.⁸⁹ Moral constructivists identify justification with truth—they believe our evidence for things constitutes whatever truth things have. Evidence *is* truth; therefore, truth is not absolute but relative to or constructed from the evidence for it.⁹⁰ Moral skeptics believe that moral truth does not exist, and therefore, they consider justification to be meaningless. Moral skeptics consider

⁸⁵ CRISPIN WRIGHT, *TRUTH AND OBJECTIVITY* 1–2 (1992).

⁸⁶ Some theories differentiate the degree of certainty required for a definition of realism or antirealism. *See, e.g.,* GOLDMAN, *supra* note 30, at 29.

⁸⁷ This is of course a simplified definition. For a detailed discussion of justification see DANCY, *supra* note 31, at 53–140.

⁸⁸ BRINK, *supra* note 4, at 126–27.

⁸⁹ *Id.* at 31.

⁹⁰ *Id.* at 31–36.

beliefs to be equally unprovable whether they are justified in some way or not.⁹¹

The question of justification is central to my criticism of Rawls's idea of "public reason." Rawls's concept is that public reason, roughly speaking, consists of those shared principles that fall within Rawls's overlapping consensus.⁹² Rawls's idea of public reason is that the shared principles are accessible to all reasonable persons and therefore can be agreed on by all reasonable persons.⁹³

What is reasonable? One definition of "reasonable" is that things that are true are reasonable. However, in Part II, it was shown that one cannot determine which concepts are true and which concepts are false when one is dealing with concepts derived by intuition, coherentism, constructivism, or skepticism. If truth is relied upon to identify public reason, a series of irreconcilable conflicts would result because any person's opinion about any subject could not be disproved, and thus any idea would have to be included in the overlapping consensus. If public reasons are selected on the grounds of truth, therefore, one will have to admit to the overlapping consensus all the ideas that are thought of as illiberal, from Nazism to slavery, if even one person holds these ideas.

Intuitive ideas are not any more unreasonable than coherentist, constructivist, or skeptical ideas.⁹⁴ For example, a person may have a belief that they are the reincarnation of a person who lived several hundred years ago. Given our lack of knowledge about what constitutes individual identity, this is not an unreasonable idea, even though it probably could not be derived from any coherentist or constructivist theory. Similarly, it is not unreasonable to believe that the Earth is only 6,000 years old and that a God created the evidence of a greater age as a sort of illusion. Many of our ideas that are based on tradition seem to have been originally derived from intuition.⁹⁵

If intuitive ideas are considered unreasonable, it seems that they would have to be considered untrue, and as has been shown above, intuitive ideas cannot be proved to be less true than other ideas. Therefore, the definition of public reason as containing only ideas that are reasonable is very problematic,

⁹¹ GOLDMAN, *supra* note 30, at 29.

⁹² RAWLS, *supra* note 9, at 212-54.

⁹³ *Id.* at 241-44.

⁹⁴ See Nagel, *supra* note 2, at 230.

⁹⁵ Rawls himself relies on intuition as the method of identifying the basic characteristics of the political enterprise. RAWLS, *supra* note 9, at 8-9, 13-14, 26, 39-40, 45, 116-18, 150-51.

at least for liberalism.⁹⁶ Rawls himself has stated that ideas outside the overlapping consensus should not be taken to be “false,”⁹⁷ but it is difficult to see how ideas outside the overlapping consensus can be taken as other than false if they are defined as unreasonable.⁹⁸ I propose that what properly defines public reason is neither truth nor reasonableness, but instead a certain type of justification—wide-circle justification—at the metaethical level.

There are two types of justification used by the three metaethical theories of intuitionism, coherentism, and constructivism. Coherentism and constructivism both use a type of circular justification.⁹⁹ Intuitionism uses self-justification.¹⁰⁰ These types of justification will be addressed in Parts III-A and III-B. The question of justification as the criterion for public reason will be re-examined in Part III-C.

A. Circular Justification

Justification is circular when a group of beliefs justify each other. For example, I might say: “I want to run with my dog every day because I want my dog to have exercise, and I want my dog to have exercise because I want my dog to be healthy, and I want my dog to be healthy so that she can run with me every day.”

Circular justification seems to be clearly invalid. It rests on nothing, and anything could be proved by circular justification.¹⁰¹ Therefore, it would seem that the fact that coherentism and constructivism rest on circular justification would be enough to invalidate them on the grounds of conceptually weak justification. However, identifying justification as “circular” does not take into account the important variable of how large or small the justificatory circle becomes.¹⁰²

In Part II, coherentism and constructivism were defined as theories in which the way beliefs fit together is important. Coherentism as a general

⁹⁶ Another serious objection to the use of “reasonableness” as a criterion for anything is simply the vagueness of the term. As Joshua Cohen says: “[t]he charge of unreasonableness is commonly a ponderous way to express simple disagreement.” Joshua Cohen, *Moral Pluralism and Political Consensus*, in *THE IDEA OF DEMOCRACY* 270, 286 (David Copp et al. eds., 1993).

⁹⁷ RAWLS, *supra* note 9, at 138.

⁹⁸ Edward B. Foley, *Rawlsian Liberalism and Establishment Clause Jurisprudence*, 43 CASE W. RES. L. REV. 963 (1993).

⁹⁹ BRINK, *supra* note 4, at 122–25.

¹⁰⁰ MACKIE, *supra* note 27, at 38–39.

¹⁰¹ BRINK, *supra* note 4, at 105.

¹⁰² For similar views, see *id.* at 122–25 and GOLDMAN, *supra* note 6, at 204–13.

metaphysical (not specifically metaethical) theory has been used to provide a basis for science and observational facts similar to that formerly provided by empirical foundationalism.¹⁰³ The difference is that, instead of basic observational facts being taken as true, the observational facts are required to form part of a coherent whole before they are accepted.

This idea of coherence as providing evidence for truth or justification relates to the discussion in Part II about the skeptical argument from error. Although coherentist theories of truth do not generally claim to provide foundational truth, they maintain that the fact of coherence itself is evidence for objective truth.¹⁰⁴ The probability of error is less in a system where all observational facts are explained versus a system where only some observational facts are explained.¹⁰⁵ The probability of error in a coherentist system may never actually reach zero, but it will approach zero more and more closely as more and more facts are fitted in a coherent way into the theory.¹⁰⁶ In a similar way, if a theory that is constructivist (and therefore does not claim ties to truth) nevertheless takes account of as many observations as possible, the probability of error of that theory will decrease as the number of observational facts included increases. Thus, a large circle of justification is conceptually stronger as an indication of truth than a small circle.

Note that it is possible for a constructivist to use a small circle of justification. This would be a type of anti-realist intuitionism. Although constructivist theories do not assume a large circle of justification in the way that coherence theories do, most constructivists attempt to derive theories that have a large circle of justification.¹⁰⁷ The reason for this is that a small-circle constructivist theory, because it does not claim any ties to objective truth in the way that intuitionism does, would be unreasonable to the extent that its justificatory circle is small. Intuitionism is not unreasonable in spite of its small justificatory circle precisely because intuitionism makes a claim to perception of objective truth, and this claim stands up to the skeptical arguments as well as any other truth claims. A small-circle constructivist theory claims neither truth nor justification, and thus is unreasonable in the sense that it is completely contingent.

What has been said here about circular justification could apply to either metaphysical or metaethical theories. However, an additional problem exists as

¹⁰³ RESCHER, *supra* note 28, at 5–40.

¹⁰⁴ BRINK, *supra* note 4, at 125–30.

¹⁰⁵ See, e.g., HILARY PUTNAM, MEANING AND THE MORAL SCIENCES 125–26 (1978).

¹⁰⁶ The obvious mathematical analog is the calculus of limits. This is not to imply that we could ever achieve a completely coherent system. See O'HEAR, *supra* note 58, at 124–36 for a summary of the argument that we can never achieve a fully coherent system.

¹⁰⁷ See, e.g., MACKIE, *supra* note 27.

to how metaethical coherence and constructivist theories fit in with metaphysical coherentist and constructivist theories. To the extent that a metaethical theory fits in with a metaphysical theory, it will have a larger justificatory circle, and thus will be stronger.¹⁰⁸ To the extent that a metaethical theory is consistent only within its moral scheme, and not within any more comprehensive theory, the metaethical theory will be weaker.

The conclusion to be drawn from this is that for a metaethical theory to have the strongest possible justification, it must have the largest possible circle of justification. Therefore, to the extent that coherentist or constructivist metaethical theories are part of a large justificatory circle, they are conceptually strong and can be part of the overlapping consensus.

B. *Self-Justification*

Self-justification can be seen as the limiting case of circular justification that is, when the circle of justification is so small that it has only one element—reference to the thing itself for its own truth. This is the type of justification used by intuitionism, which derives truth from self-reference. In the previous discussion on intuitionism, it was mentioned that the intuitionist believes that humans have a special faculty for detecting moral truths. The self-justification in intuitionism is that the intuitionist believes something is true because her special faculty tells her it is true, and the special faculty determines truth because it is the nature of the special faculty to do so.

The validity that intuitionism possesses comes from the fact that one cannot disprove intuitive ideas in the sense that one cannot say they are false. However, it can be said that intuitive ideas are poorly justified. As noted above, anything can be proved by circular justification. Intuitionism, however, does not have the benefit of large-circle justification which at least partially saves coherentism and constructivism from this objection.

Intuitionist ideas are often thought of as religious ideas, and the extent to which government should or should not adopt religious ideas has generated a huge literature.¹⁰⁹ However, religious ideas are far from the only intuitionist ideas that have affected governmental policies.¹¹⁰ Two other categories of

¹⁰⁸ BRINK, *supra* note 4, at 127.

¹⁰⁹ See, for example, works by Kent Greenawalt, Thomas Jefferson, James Madison, Michael McConnell, Michael Sandel, and Steven Smith.

¹¹⁰ As Jean Hampton says: "when values are free-standing, our fundamental commitments are to the values themselves, and not to the theoretical speculations we generate to explain them." She also notes that Thomas Jefferson said of the values on which our country was founded: "We hold these truths to be self-evident." Jean Hampton, *The*

intuitionist ideas that are often acted on by government are ideas derived from common sense and ideas derived from tradition.

Einstein once referred to common sense as the layer of prejudices laid down in our minds before we reach the age of eighteen. Perhaps the best demonstration that many of our common-sense ideas are false is the theory of quantum physics.¹¹¹ But in addition to conundrums from abstruse areas of science, there are more ordinary ideas that many people accept based on common sense. For example, people once accepted as common sense the idea that the earth was flat.

Charles Larmore argues that it is legitimate to use tradition as a justification for morality.¹¹² William Galston also argues for tradition as a basis for morality in government.¹¹³ Communitarians such as Michael Sandel have argued that we are defined in part by our traditions.¹¹⁴ I contend that a belief derived from tradition is an intuitionist belief. One might argue, against this notion, that tradition is not taken to be self-evident, as intuitionist ideas are, and is therefore not intuitionist. One could say that traditional rules are those that have proven themselves over time or that humans need traditional rules to function. If tradition is justified by some other means, such as utility, it is actually utility and not tradition that is being justified. But if a tradition is adopted solely because it is part of the body of tradition, then it is being justified by itself. That is, the justification of a tradition is on an intuitional basis.

David Brink makes an argument based on tradition in his defense of moral realism in his book *Moral Realism and the Foundations of Ethics*.¹¹⁵ The important aspect of Brink's ideas about tradition in regard to the thesis proposed here is not that he uses tradition to support the concept of moral realism but that he argues tradition should be accorded weight merely by the fact of its existence. Brink writes that, because we have a tradition of talking and acting as though moral beliefs are real facts, anyone who claims otherwise should have to prove her claim. Brink refers to this as placing the "burden of

Moral Commitments of Liberalism, in *THE IDEA OF DEMOCRACY*, *supra* note 96, at 292, 302-03.

¹¹¹ As Niels Bohr said, "If you are not mystified by quantum theory, you haven't understood it."

¹¹² Charles Larmore, *Beyond Religion and Enlightenment*, 30 *SAN DIEGO L. REV.* (forthcoming 1994).

¹¹³ GALSTON, *supra* note 17.

¹¹⁴ Michael J. Sandel, *Freedom of Conscience or Freedom of Choice?* in *ARTICLES OF FAITH, ARTICLES OF PEACE* 74 (James D. Hunter & Os Guinness eds., 1990).

¹¹⁵ BRINK, *supra* note 4.

proof" on the opposition to these traditions.¹¹⁶ Brink supports this presumption by saying that these traditions are our natural starting point, consisting of common sense, our language, and our actions.¹¹⁷ However, at least three objections exist to using tradition as evidence for the truth of any proposition.

The first objection is simply that tradition may not be very helpful for governmental policy. For example, humans have a tradition that murder is wrong. This is a tradition that virtually all members of society agree upon.¹¹⁸ However, the very fact that this moral tradition is so widely accepted means that a government in a liberal society can penalize murder without interfering with the beliefs of any substantial segment of that society.¹¹⁹ Remember that the liberal paradox is the problem of enforcing moral claims on those who disagree with the validity of those claims. If everyone agrees that murder is wrong, no problem exists for liberalism.¹²⁰ The situation is different with controversial questions like abortion and animal rights.

Thus, if tradition is seen as deriving its legitimacy from unanimity, it is not so much wrong as just not very useful. However, if we allow tradition to include concepts about which there is no unanimity or even virtual unanimity, we have a new problem. To the extent that a tradition does not command unanimity, it is a double-edged sword in that tradition can be invoked to prove virtually any controversial point. A clear example of this is the discussion of the tradition of society in regard to homosexuality in *Bowers v. Hardwick*.¹²¹ In that case, the majority stated that the tradition in American society was that homosexuality was unacceptable, while the dissent questioned how comprehensive this "tradition" was.

The second objection to the use of tradition as a justification for governmental policy is that tradition changes. If tradition is used as the arbiter of values, it must either be admitted that tradition is fallible or that values change over time. If we make either of these admissions, it adds to the complexity of how we should attempt to use tradition. Should we try to determine the direction of change in tradition and then consciously move in that direction? Or should we try to resist any change as against the nature of tradition?

¹¹⁶ *Id.* at 35–36.

¹¹⁷ *Id.* at 23–31.

¹¹⁸ There is an important distinction for liberalism between true unanimity and virtual unanimity.

¹¹⁹ See Solum, *supra* note 14 ("When everyone accepts the premises and inferences that justify a policy, then public justification of the policy is least urgent.").

¹²⁰ Some theorists believe that even unanimity is not a justification for governmental action because people may change their minds, or people not yet born may disagree.

¹²¹ 478 U.S. 186, *reh'g denied*, 478 U.S. 1039 (1986).

With animal rights, for example, what tradition would we look to? If the question the government has to decide is whether to allow factory farming, does the government look to the tradition of the last 20 years when factory farming has become the norm? Or is the tradition the more humane practices of farmers in this country from 1800 to 1940? Or should we go even further back to the times when humans did not have domesticated animals?

The third objection to the use of tradition as a justification for governmental policy is simply that it begs the question of justification. Accepting tradition as a guide, without more, simply means defining justification as what accords with tradition. This is self-justification, which raises the same objection as the objection to intuitionism. One could argue that Brink does not blindly accept tradition as justified since he says that it is merely a presumption of justification.¹²² In an area as contentious as governmental enforcement of morality, however, a presumption can be a very powerful starting point.

Larmore points out that we are creatures of tradition and that rejecting our tradition would be foolish if not impossible.¹²³ However, even though we may choose to be guided by tradition in our private lives, or even if we are compelled by our nature as human beings to be guided by tradition in our private lives, should tradition be a consideration in the public sphere? For example, what if there were clear evidence that a certain tradition, such as our society's pattern of distribution of wealth, appeals to us because of instincts we developed during the times when humans lived in small tribal societies? Even if we find it impossible to totally go against our instincts in our private lives, it is questionable whether government should adopt instinct-based tradition as a presumption of correct public policy.

The problem with using tradition as a basis for governmental action mirrors the problem of using intuitionism as a metaethical theory. The justificatory circle is very small. As discussed in Part III-A, the smaller the justificatory circle, the higher the probability of error. Thus, because intuitionism by definition justifies beliefs by a small justificatory circle, intuitionism is a conceptually weak theory for government to use. Because of this conceptual weakness, intuitionism should be excluded from the overlapping consensus.

¹²² BRINK, *supra* note 4, at 31.

¹²³ Larmore, *supra* note 112.

C. Justification as a Basis for Public Reason

This section returns to the question of defining public reason, and why wide-circle justification, not reasonableness, should be the criterion for selection of public reason. Rawls defines public reason in a somewhat different way than will be done here. Rawls sees public reason as derived by the people and accepted by each person based on their individual moral code. Thus, Rawls believes government can be metaethically neutral because it is the people, not the government, selecting public reasons.¹²⁴ In contrast, this Comment sees public reason as derived by the government and accepted by the people. Thus, government cannot be metaethically neutral.

There are two actors in a governmental system. One actor is the government itself, and the other actor is the governed people. The concept of public reason applies to these two actors in different ways. Individuals can accept shared public reasons based on each individual's comprehensive moral code. This personal moral code may contain ideas derived from intuition, from tradition, from a coherentist philosophy, or from any other metaphysical or metaethical ground. Thus, public reason for individuals is directly related to their individual comprehensive moral code. But before an individual can accept or reject a public reason, the government must put forward a proposal. For example, if government does not offer free public education, the question of whether creationism should be taught in the public schools never arises. If government does not propose to regulate abortions, the question of abortion has nothing to do with public reason.

Granted, a dissatisfied populace can stage a revolution or otherwise modify a government if the government does not respond to the people, but at any given point in time it is the government, not the people, that must select actions that correspond with public reason from all of the actions available to the government. A moment exists when a policy is selected. In our system, there are several crucial moments, including proposal of a policy by the executive, votes in the legislative process, and presentment. But at each crucial moment in the selection or rejection of a policy, it is the government that is making the choice and not the people. The people make long-term and general choices, but the people delegate their power to make the day-to-day choices to government.

One could analogize to a spendthrift trust, where the government is the trustee and the people are the beneficiary. In such a trust there are two participants, and there is a pre-arranged structure that regulates their

¹²⁴ See, e.g., Solum, *supra* note 14 ("Rousseau's belief that there is a clearly demarcated divide between public reason and private interest is not a feature of Rawls's idea.").

relationship. The beneficiary could perhaps overpower the trustee and take the corpus of the trust by force, which would be analogous to a revolution. But outside of that event, the trustee has its own identity separate from the beneficiary.

An analogy that might be closer to the relationship between government and the people in a liberal democracy is the analogy between a corporation and its stockholders. The stockholders initially choose the way the corporation will be run, and they retain the right to modify the corporation, but the corporation takes on its own identity. Indeed, corporations become legal *persons* in many respects. If we view the government and the people in a liberal democracy as essentially the same, as Rawls does, then we lose sight of an important distinction, which is that public choices are made by an entity that is created by the people but not identical with the people. Collapsing this distinction between the government and the people is part of what leads to difficulties with the liberal paradox. This notion will be defended in the remainder of this Part.

If we assume that it is government that selects public reasons and that a government does not have a personal moral code, by what method can a government select public reasons? The difference between a government and an individual is that a government orders the relationships between people. Thus, we must look at what the difference is between an individual's moral code and a moral code enforced by government in light of the fact that the function of government is defining structures for interpersonal relationships.

In looking at the justification of metaethical theories, this Comment concluded that intuitionism uses a very small circle of justification. In fact, intuitionism uses the smallest possible circle of justification, which is a thing justifying itself by itself. Intuitionism is a uniquely personal metaethical theory since the circle of justification does not reach beyond the individual person. Because there is no completely satisfactory answer to the skeptical argument from error, one could argue that an individual can make arbitrary moral choices or no choices at all. The individual could choose to ignore either truth or justification, or both, or to give them little weight, or to define her own justification because individuals are free to adopt their own personal moral codes, within the limits of law.¹²⁵

Government, on the other hand, should strive for the widest possible circle of justification because government is a uniquely public activity. Government is not just one individual, but rather a relationship of all individuals, and thus government should not use a justificatory scheme such as intuitionism which is merely personal. Government must use the widest possible justificatory scheme

¹²⁵ I do not mean to imply that it is morally correct for persons to derive their own moral code, only that it is possible within the limits of the law.

because government is the widest possible relationship—that of all people to all other people. In other words, government must strive to be defined by impartiality. Thomas Nagel says, “If liberalism is to be defended as a higher-order theory rather than just another sectarian doctrine, it must be shown to result from an interpretation of impartiality itself, rather than from a particular conception of the good that is to be made impartially available.”¹²⁶

Thus, the idea of the impartiality of government is derived, not by intuition as in Rawls’s theory, but from an examination of the nature of government. Impartiality of the government arises from the notion of the government as a separate entity from the people that is devised for the purpose of regulating relationships between people. To be impartial, the government must adopt justifications for its actions that apply to all of the governed people.

Therefore, the fundamental criterion of public reason should be wide-circle justification. The larger the justificatory circle, the more accessible the public reason will be to people in general; thus, the fundamental criterion of public reason can be restated as “accessibility.”¹²⁷ Although government and individuals may be equally concerned with the *truth* of a moral code, government must be more concerned than the individual need be with the *justification* of a moral code.¹²⁸ Justification can be merely personal (intuitional) for the individual; but justification for public policies selected by government should be wide-circle.¹²⁹

An objection to the idea that government should consider the justification as well as the truth of metaethical theories is that this position is realist and thus favors realism over antirealism. That is, antirealism collapses the distinction between truth and justification whereas realism does not, and if government does not collapse this distinction, government is adopting a realist perspective. However, government that considers justification as well as truth in selecting metaethical grounds only recognizes truth in the negative sense of maintaining neutrality toward the question of what is true.

One could make a persuasive coherentist/constructivist argument that a government that is not fair and does not treat humans as free and equal persons

¹²⁶ Nagel, *supra* note 2, at 223.

¹²⁷ For a similar idea, see Michael J. Perry, *Toward an Ecumenical Politics*, 60 GEO. WASH. L. REV. 599 (1992).

¹²⁸ This conception of public reason is similar to that of Rousseau. See Solum, *supra* note 14.

¹²⁹ Note that this idea does not presuppose the desirability of a liberal government. I would argue that a liberal government is most in accord with a comprehensive coherentist or constructivist theory, but in this scheme, the desirable type of government is derived from metaethical theories, not the other way around.

is a government that is inherently unstable.¹³⁰ Thus, wide-circle justification may well derive a starting point that is similar to Rawls's intuitive starting point, but it is derived by means of a metaethical argument that bases the requirement for freedom and equality on the government's nature as an arbiter between people.

To say that a government must be impartial and must consider justification is not the same as saying that a government must treat all people in the same way. For example, it might be decided on the basis of a well-justified argument that in order to treat all people impartially, the government must make allowances for the fact that not all people are similarly situated. Or, the government might decide to adopt a program such as Head Start on the grounds that to treat some children impartially they must be treated differently.

To say that a government must consider justification is also not to say that a government must be majoritarian. Accessibility and public reason as defined by justification have nothing to do with majoritarianism. For example, it is quite possible that 95% of the members of a given population might have the same idea derived by each individual from their own intuition. If this idea (for example, that the world would end in three years) was not derivable from any well-justified coherentist or constructivist theory, it would not be part of public reason. Of course, an intuitive belief held by a large percentage of the population might well affect government action, but it will do so by brute

¹³⁰ The argument is as follows: Social action involves paradoxes such as the prisoner's dilemma and free-riding. One way of looking at the strong moral feelings that people have (i.e. the strong intuitive appeal of concepts like the moral equality of all persons) is that we have psychological imperatives built-in to cooperate in social actions. This enables humans to arrive at the most efficient solutions to the paradoxes of social cooperation. See O'HEAR, *supra* note 58, at 266–68. Although we might realize intellectually that we could free-ride in a particular situation, and although that might be the logical thing to do for the individual, we don't do it because we would feel guilty. These widespread (although not universal—natural selection is not perfect) instinctive feelings allow humans to arrive at a degree of social cooperation that benefits us all without allowing individuals to ruin the scheme through too many solely self-interested actions. It can be argued that it is therefore perfectly reasonable for government to act in a way that recognizes humans' instinctive need for social cooperation, thus entailing that people must be considered to be to some extent equal or have equal rights. As Thomas Nagel says:

The requirement of impartiality can take various forms, but it usually involves treating or counting everyone equally in some respect—according them all the same rights, or counting their good or their welfare or some aspect of it the same in determining what would be a desirable result or a permissible course of action.

Nagel, *supra* note 2, at 215.

force, as it were, not through public reason.¹³¹ To the extent that intuitive majoritarian ideas do influence government, they represent a failure of public reason.

IV. METAETHICAL NEUTRALITY AND THE OVERLAPPING CONSENSUS

In this Part, the connection between the various ideas that have been discussed so far—metaethical theories, truth, justification, public reason, Rawls's ideas of governmental metaethical neutrality, the overlapping consensus, and the constructive definition of the person—will be examined. A new overlapping consensus that incorporates the ideas of wide-circle justification and a concept of public reason that is based on accessibility will be defined.

Section A of Part IV describes the Rawlsian overlapping consensus and Rawls's definition of the person. Section B of Part IV examines an overlapping consensus that is not metaethically neutral and concludes that such an overlapping consensus does not require a problematic definition of the person to warrant government enforcement of moral rules. Section C of Part IV discusses some of the problems with applications of metaethical theories.

A. Rawlsian Liberalism and its Relation to Metaethics

Rawls's theory of justice is wide-ranging and general. This is good both because the broad concepts he has developed are important and because the room for interpretation that Rawls leaves has stimulated much useful discussion. The other side of the coin, however, is that Rawls's theory can be interpreted in many different ways. Therefore the person who attempts an interpretation is engaged in a dangerous task. The following interpretation of Rawls's work is put forward as only one of several possible, reasonable ways of reading Rawls.

Rawlsian liberalism contains three important related ideas which will now be evaluated as a unit. The first idea is that a conception of the political sphere should properly be neutral between various metaethical theories. The second idea is that an overlapping consensus can be derived that will provide sufficient shared public reasons for governmental decisions. The third idea is the constructive definition of the person. The relationship between these ideas is Rawls's concept that the public reasons which make up the overlapping consensus can be accepted by all citizens by reference to their individual

¹³¹ The classic discussion of majoritarian influence on government is JOHN H. ELY, *DEMOCRACY AND DISTRUST* (1980).

comprehensive moral theories, without government having to endorse any of those theories,¹³² if "all citizens" is understood to mean reasonable citizens operating according to Rawls's constructive definition of the person.¹³³

The idea behind metaethical neutrality is that people of differing metaethical views should be able to come to a consensus on the actions that *government* should take, whether they can all agree on a comprehensive moral code or not. This is the central idea of the overlapping consensus.¹³⁴ The public reasons that make up the overlapping consensus are derived by the device of the original position,¹³⁵ in which people choose what sort of society they would want if they did not know what their position in the society would be. In other words, they are behind a "veil of ignorance."¹³⁶

As Kent Greenawalt and others have pointed out, the person in the original position must have some characteristics in order to carry out the job of selection and decisionmaking.¹³⁷ That is, the person might be seen as a religious believer who believes that women should not take part in public life, or the person might be a humanist, or a utilitarian. Rawls has tried to answer this objection by stating that, while he does have a "definition of the person" as the person who makes the selection in the original position, the definition is a representation only.¹³⁸ Thus, the definition of the person is merely a device to allow the principles of public reason to be selected.¹³⁹ Rawls defines his concept of the person by beginning with the idea that people are "free and equal."¹⁴⁰ From this idea, Rawls derives several other characteristics of the person, sufficient to enable such a representational person in the original position to select public reasons that can be part of the overlapping consensus.¹⁴¹

Before looking at the substance of these aspects of Rawls's theory, it is helpful to examine Rawls's use of the idea of a "representation." The use of a representation in order to simplify an inquiry is a legitimate way of proceeding only if the representation follows certain rules. One type of representation is a

¹³² RAWLS, *supra* note 9, at 12-13, 144-45.

¹³³ *Id.* at 36-38, 144.

¹³⁴ *Id.* at 150-51.

¹³⁵ *Id.* at 22-28.

¹³⁶ *Id.* at 23.

¹³⁷ GREENAWALT, *supra* note 7, at 51-54.

¹³⁸ RAWLS, *supra* note 9, at 25.

¹³⁹ *Id.* at 93.

¹⁴⁰ For a criticism of Rawls's idea of persons as "free and equal" as an insufficient basis for deriving principles of justice, see THOMAS NAGEL, *EQUALITY AND PARTIALITY* (1991).

¹⁴¹ RAWLS, *supra* note 9, at 47-88.

simplified picture such as a map, and, like a map, the representation must reflect what it is representing. A related use of a representation is to solve a problem in a situation when there are so many variables that finding a solution is too difficult. In order to use a representation, the essential variables must be preserved and inessential variables removed. Then, when a solution has been reached, the inessential variables can be added back into the equation. Rawls's representation, however, does not claim any connection to the underlying situation he is trying to solve.

Rawls tries to simplify the problem of choosing a system of justice by defining a representational person who possesses certain characteristics which Rawls justifies on the grounds that they are intuitive. This is somewhat like solving the problem of abortion by defining human beings as not including fetuses. Rawls's definition of the person does simplify the problem, but it does not solve it because Rawls has removed some essential variables. Rawls tries to simplify the problem of selecting a just system by defining a particular type of person to make the selection. The problem with this is that there is no way for Rawls to add the variables back in. That is, there are many types of people as well as many essential human characteristics that Rawls leaves out of his representational scheme and never adds back in.¹⁴²

The substance of Rawls's scheme will now be examined. As stated above, a definition of people as "free and equal" could also be derived from a wide-circle justification scheme. So, my criticism is not so much with Rawls's starting definition of people as free and equal (although, as noted above, I would criticize the way he arrives at that definition) as with the way Rawls expands the concept of people as free and equal into a full-blown conception of the person derived by intuition.

Rawls cannot stop with just free and equal in the definition of the person since such a conception would be too thin as a basis for governmental policy.¹⁴³ In order to provide an idea of the person that would be sufficient for choices made behind the veil of ignorance, Rawls expands his conception of the

¹⁴² This is not to say that Rawls's construction of political philosophy as a subset of philosophy in general is invalid. Rawls's idea is that the construction that is involved is a framework or a subunit, if you will, of a comprehensive doctrine. The political sphere is the shared part of each person's comprehensive doctrine, and all the other parts of the doctrine can be developed as the individual sees fit. Recognizing the difference between the political arena and a comprehensive arena is different from making a representation.

¹⁴³ Freedom and equality can be seen as in conflict. The most obvious example is in the distribution of wealth. Because of differences in human abilities, complete freedom will lead to humans who are more talented at amassing wealth accumulating a higher than average percentage of overall wealth. Obtaining equality necessarily means decreasing the freedom of the most talented.

person to include characteristics that are considerably more controversial and questionable than free and equal.

Rawls includes in his notion of the representation of the person the idea that the "person" must be capable of reciprocity.¹⁴⁴ Rawls's notion of reciprocity is that people in society gain the right to participate in society, including its protections and obligations, by being creatures who can reason with others in a way that can include reciprocal recognition of obligations. In other words, persons are not just free and equal; they must also have the ability to contract. Not only must they have the ability to contract, but they must also use this ability to actually engage in reciprocity, or they are unreasonable. Rawls thus defines reasonableness to include an element of reciprocity.¹⁴⁵

This definition of reasonableness as including reciprocity is very important in Rawls's scheme because it enables him to say that people who will not engage in a dialogue of give and take are not reasonable. Because Rawls has defined public reason as those ideas that are reasonable, this means Rawls can exclude the ideas of those who will not or can not engage in compromise ("reciprocity") from the overlapping consensus.¹⁴⁶

This is Rawls's solution to the paradox of liberalism. By using a constructive definition of the person that includes a notion of reciprocity supposedly derived from the notions of freedom and equality, Rawls has devised a way to exclude intolerant ideas and intolerant people from the overlapping consensus. Intolerant people, meaning those who will not engage in reciprocity, cannot select concepts of social justice behind the veil of ignorance.

One might not disagree with the exclusion of intolerant ideas from the overlapping consensus and yet still believe that the constructive definition of the person is an indefensible way to achieve this result. The constructive definition of the person leads to far more problems than it solves. This definition of the person leaves Rawls's theory open to at least four serious objections. The first and perhaps most serious objection is that, with no reference to an underlying metaethical theory, the definition of the person is contingent. Why should we accept the notion of rights-holders defined as free and equal persons on a contingent basis, much less defined as only those who have the capacity to contract?¹⁴⁷ If we are going to accept contingent definitions, why should we not accept definitions of humans as altruistic, or of humans as solely pleasure-seeking, or of humans as a blight on the earth? Ideas

¹⁴⁴ RAWLS, *supra* note 9, at 48-54.

¹⁴⁵ *Id.* at 16.

¹⁴⁶ *Id.* at 195-97.

¹⁴⁷ Hampton, *supra* note 110, at 299-300.

that are justified only by their intuitive appeal beg the question of truth. As Jean Hampton says:

Often in his recent work, Rawls stresses the fact that the content for the sort of overlapping consensus he is commending depends on the political society's being a modern Western democracy with certain fundamental beliefs (e.g., in freedom and equality) that are part of the intellectual culture of that society. . . .

Taking such a position, however, will undermine the effectiveness of his defense of his theory of justice in the eyes of many philosophers. On this view Rawls would be arguing: "If a society believes in x, y, and z, then the theory of justice I commend will be shared by all the members."¹⁴⁸

The second objection to Rawls's conception of the person is that the definition of reasonableness to include the notion of reciprocity seems quite strained. Reasonableness and reciprocity are related, but reasonableness as normally understood does not include a notion of reciprocity as a necessary element. For example, it would not be unreasonable for a government to adopt a system of absolute power vested in philosopher-kings, but such a system need not involve reciprocity. Any merit-based selection system would be reasonable insofar as the selection criteria is related to the job to be done, but to the extent that a system is merit-based, it is not reciprocal.

Rawls's concept of "reasonableness" is both overinclusive and underinclusive as a basis for public reason.¹⁴⁹ It is overinclusive because it could include things based on intuition such as tradition and common sense which are not necessarily unreasonable but do not necessarily have wide-circle justification.¹⁵⁰ It is underinclusive because Rawls's notion of reciprocity as part of reasonableness brings in an element of majoritarianism that could lead to rejection of ideas based on wide-circle justification merely because a majority does not approve of the ideas.

The third objection to the Rawlsian conception of the person is the definition of the person as one who has an ability to contract. This definition leads to the problem of how Rawlsian liberalism can encompass the "borderlines of status" questions¹⁵¹ such as mentally disabled humans, very young children, fetuses, and animals, all of whom arguably cannot enter into the type of reciprocal relationships with normal adult humans that Rawls envisions as part of the definition of the person. If governmental protection is

¹⁴⁸ *Id.*

¹⁴⁹ *Id.* at 306–09.

¹⁵⁰ See *supra* note 95 and accompanying text.

¹⁵¹ GREENAWALT, *supra* note 7, at 98–172.

completely dependent on a type of contractual ability defined as something which only normal, adult humans have, creatures of borderline status are arbitrarily defined out of the moral sphere of those creatures who can have rights.¹⁵²

The fourth problem with the Rawlsian conception of the person is that because the Rawlsian definition of the person is contingent, one cannot even dispute his conception of the person in the way that one could dispute the conceptual strength of a metaethical theory. If government bases governmental justification for moral rules on a public debate that has included norms derived only from wide-circle justification, it is then possible for a public debate to take place on the question of whether government has included the correct metaethical theories in an overlapping consensus. In other words, the contents of the overlapping consensus itself are still open to question and can change over time. Rawls's definition of the person, on the other hand, because it is contingent, is immune to attack in Rawls's scheme on grounds of its conceptual strength or weakness. Because the idea of the person is a definitional construction, it is beyond question.

The supporter of the Rawlsian definition of the person might argue that this definition of the person follows the logic of Kant, in that, rather than trying to justify a conception of the person on first principles, it merely recognizes what in fact exists.¹⁵³ This argument says that rather than try to define reality or trace it to its roots we should start from the point where we are now and merely describe reality as it is.

The problems with a Kantian definition of the person are many. This argument does not adequately answer the objection that a Kantian description of the person merely begs the question of truth. Furthermore, attempting to describe the nature of humans gives a great deal of discretion to the describer. One could make quite a good argument that selfishness and violence are just as much a part of human nature as the concepts of freedom and equality. And, even if we do succeed in describing what exists (what *is*), this does not mean that what exists is what *ought* to exist.¹⁵⁴ Finally, a conception of the person that is contingent and beyond challenge does not provide a flexible enough basis for government to change over time.

¹⁵² See, e.g., PETER CARRUTHERS, *THE ANIMALS ISSUE: MORAL THEORY IN PRACTICE* 98-121 (1992) ("Animals will, therefore, have no moral standing under Rawlsian contractualism, in so far as they do not count as rational agents.").

¹⁵³ RAWLS, *supra* note 9, at 99-100.

¹⁵⁴ DAVID HUME, *An Inquiry Concerning Human Understanding*, in *ENQUIRIES CONCERNING HUMAN UNDERSTANDING AND CONCERNING THE PRINCIPLES OF MORALS* 5 (L.A. Selby-Bigge ed., 3d ed. 1975).

The problem of how government can justify enforcing a moral code on those who disagree is a complicated problem. But Rawls's solution, which is to give up the attempt to evaluate the conceptual strength of metaethical theories and instead adopt a contingent definition of the person, creates more problems than it solves.

Rawls's idea of wide reflective equilibrium, if interpreted as a requirement for wide-circle justification, would negate the need for a complicated, contingent definition of the person. This idea will be developed further in the next section. Rawls apparently does not see wide reflective equilibrium as requiring wide-circle justification since he defines public reason as including conclusions derived from "commonsense knowledge"¹⁵⁵ and since he bases his entire theory on intuitive starting conceptions.¹⁵⁶

B. *A New Overlapping Consensus*

Rawls's attempt to bypass metaethical disagreements (as interpreted here) founders on the necessity for a government that is metaethically neutral to make decisions about public reason based on a contingent definition of the person. The benefit of metaethical neutrality, that hard questions of truth and justification do not have to be addressed, is not worth the cost of excluding illiberal ideas on a contingent basis.

Rather, the goal of excluding illiberal ideas from the overlapping consensus can be achieved without resorting to a contingent definition of the person. An evaluation of the truth and the justification of metaethical theories, with exclusion of theories with weak claims to truth or to wide-circle justification, will result in illiberal ideas being excluded from the overlapping consensus on a conceptually stronger and more consistent basis than the contingent definition of the person.

Most people, if asked to describe particular illiberal ideas that should be excluded from the overlapping consensus, would name examples such as slavery, government prescription of religious belief, lack of due process, and so on. Rawls takes the view that we arrive at our definition of illiberal ideas in an intuitive way.¹⁵⁷ Rawls says we cannot justify the exclusion of illiberal ideas from the overlapping consensus on any metaethical basis because we cannot prove the truth of any metaethical theory. We cannot prove that slavery is right or wrong, Rawls would say, by appeal to any metaethical theory that all citizens would accept because we cannot prove that the citizens who do not

¹⁵⁵ RAWLS, *supra* note 9, at 139.

¹⁵⁶ See *supra* note 95 and accompanying text.

¹⁵⁷ RAWLS, *supra* note 9, at 116–18, 123–25.

accept a particular metaethical theory are wrong. Therefore, our definition of illiberal ideas must be contingent.

As the discussion of the skeptical possibilities showed, one cannot determine with certainty what is true. If Rawls's definition of the person is seen as the only alternative to skeptical paralysis, perhaps the sacrifice of adopting a contingent definition of the person will seem worthwhile. However, it is overly formalistic to conclude that, because we cannot have perfect justification or perfect assurance of truth, we must therefore adopt neutrality between all forms of justification and truth.

This Comment proposes that the reason we think of certain ideas as illiberal is precisely because these ideas tend to be ideas that could only be warranted by metaethical theories such as foundationalism and intuitionism that are weak in respect to either truth or wide-circle justification. We may think of ourselves as knowing what illiberal ideas are intuitively, but in fact our conception of what is illiberal changes radically with social change. One hundred and fifty years ago, a majority of the population in certain parts of the country passionately believed that slavery was not an "illiberal" idea. One hundred years ago, most males in this country believed that females were inferior in some ways to males. One hundred years from now, it may be considered illiberal to disregard the interests of animals. The fact that our ideas about what is liberal and what is illiberal can change indicates that these ideas are not unchanging truths but are based to at least some extent on what we experience in our culture.

In Part III, the metaethical theories of coherentism and constructivism were identified as having stronger justification than foundationalism and intuitionism. This Comment now proposes that these metaethical theories can be used, through the means of public debate, to identify illiberal normative ideas. Theories that have poor justification should be excluded from the overlapping consensus as well as normative conclusions that can only be reached by means of a poorly justified theory. This approach does not require government to make a choice between realism and antirealism, but it does require government to make a choice for normative ideas that correspond to well-justified metaethical theories rather than normative ideas that correspond to poorly justified metaethical theories.

When poorly justified metaethical theories are excluded from the overlapping consensus, individual normative ideas that are derived from those theories might be excluded but will not necessarily be excluded. For example, if one person had an intuitive feeling that women should have an absolute right to terminate pregnancies any time before birth and another person had an intuitive feeling that abortion was always morally wrong, one could argue that these ideas would not automatically be considered illiberal (and thus excluded

from the overlapping consensus) just because the individuals derived them by intuition. The reason they would not automatically be excluded is because it would be possible to come to either of these ideas based on conclusions derivable from a coherentist or constructivist metaethics.

However, if the deep-rooted belief of an intuitionist was that the world will end in three years because the Bible predicts it, this belief would be excluded from the overlapping consensus. An appeal to the infallibility of the Bible has an extremely small circle of justification, and the idea that the world will end in three years does not correspond to any coherentist or constructivist theory of the lifespan of earth that is based on a large justificatory circle. The ideas derived by intuition will not therefore be excluded automatically from the overlapping consensus, but will be excluded only when they are not derivable by coherentist or constructivist theories.

The theory advanced here of a new overlapping consensus can be reconciled to some extent with Rawls's theory by recasting the idea of wide reflective equilibrium as a theory for government as well as for individuals. If the requirement for complete metaethical neutrality is abandoned, wide reflective equilibrium could be seen to include a requirement that ideas accepted by government would have to be capable of derivation from a coherentist or constructivist scheme with a wide circle of justification. The concept of wide reflective equilibrium seems to be a coherentist or constructivist concept by its definition. The view of wide reflective equilibrium expressed in this Comment is that wide reflective equilibrium is something the government uses on the basis of justification rather than something the individual uses on the basis of truth, which would be more in accord with Rawls's view.

What is the difference in what would be excluded from the overlapping consensus by Rawls's idea of reasonableness, and this Comment's view of wide-circle justification? At first blush there might not appear to be much of a difference, but in fact the difference is large. It would be a mistake to think that intuitionism no longer has a hold on political philosophy. As noted above, Larmore refers to tradition,¹⁵⁸ and Rawls refers to "fundamental ideas we seem to share through the public political culture."¹⁵⁹ In other words, Rawls relies heavily on intuitive ideas in his theory, in deriving the starting point of the theory, in defining the person, and in the selection of public reason.

If the overlapping consensus is restricted to ideas derivable from metaethical theories that use strong justification, there is no longer a necessity for a contingent definition of the person. Illiberal ideas can be excluded from

¹⁵⁸ Larmore, *supra* note 112.

¹⁵⁹ RAWLS, *supra* note 9, at 150.

the overlapping consensus based on their metaethical failings rather than on their failure to coincide with a definition of the person. This new overlapping consensus would not need to make the strained definitions that Rawls must make to support his conception of the person.

C. Problems with Applications of Coherentism and Constructivism

This Comment has used some normative examples in this paper to illustrate particular points, but on the whole normative examples have been avoided. This paper is not about normative conclusions. One of the characteristics of both Rawls's concept of the overlapping consensus and this new overlapping consensus is that the overlapping consensus can contain many conflicting normative ideas.¹⁶⁰ For example, no coherentist or comprehensively constructivist metaethical theory provides an obvious answer to the problem of abortion. The overlapping consensus might contain many contradictory normative ideas that would have to be resolved by public debate. Different normative conclusions can be derived from different conceptually strong metaethical theories. Selecting theories for how normative results are to be reached in the event of disagreement over well-justified metaethical alternatives is the function of a government's constitution, which can be viewed as an intermediate step between metaethics and laws. However, it is now appropriate to deal briefly with some general objections about the derivation of normative ideas from this new overlapping consensus.

One could argue that almost any idea could be justified within a comprehensive coherentist or constructivist scheme given our limitations of knowledge.¹⁶¹ However, in fact wide-circle justification offers a high degree of guidance. The danger of a personal agenda or prejudice exists in any area of uncertainty, not just in wide-circle justification schemes. We can only do our

¹⁶⁰ As Thomas Nagel says:

It is important to stress that the nondogmatic moral disagreements which fall within the public domain may nevertheless be irresolvable in fact. That there is common ground does not mean that people will actually reach agreement, nor does it mean that only one belief is reasonable on the evidence

The idea is that in such a case there is a common reason in which both parties share, but from which they get different results because they cannot, being limited creatures, be expected to exercise it perfectly.

Nagel, *supra* note 2, at 234.

¹⁶¹ See, e.g., STEPHEN J. GOULD, *THE MISMEASURE OF MAN* (1981).

best to make progress in the right direction, knowing that our efforts will not be perfect.

Another possible objection to an overlapping consensus that contains only widely justified coherentist or constructivist ideas is that such a consensus favors science. As a first response to this objection, it is important to note that scientific conclusions are not the only things that have wide-circle justification. An important category of things that could be included in a coherentist or constructivist philosophy would be analogy to successful systems. But to the extent that science is well-justified, objecting to scientific conclusions just because they are scientific seems nonsensical. Objections should be directed at the metaethical theories themselves and at the new overlapping consensus itself, rather than at where they lead. Otherwise, a conclusion in favor of a theory opposing science will be merely results-oriented. If wide acceptance of scientific ideas is the result of wide-circle justification, this by itself is an argument for science rather than an indictment of wide-circle justification.

In a similar vein, one could object that a wide-circle justification theory would be utilitarian. I do not believe this is the case, however, at least not in any simplistic sense. There are many ideas that could be included in the new overlapping consensus that have the importance of individual autonomy as a central tenet.¹⁶² Similarly, drawing a distinction between government action and private action negates the criticism of utilitarianism that it requires individuals to adopt the best possible action at all times and that utilitarianism thus is opposed to individual autonomy. In the new overlapping consensus, government has the duty to do good all the time, individuals do not.

This Comment has also not addressed the question of how the line between the personal and the political applies in practical terms.¹⁶³ For example, can a "good citizen"¹⁶⁴ vote based on intuitive ideas or must the good citizen vote only on the basis of ideas that can be derived within the overlapping consensus? And even if private citizens can vote based on their intuitive ideas, what about legislators? What should be the terms of public debate? Should public debate include non-public reasons?¹⁶⁵ Public reason, in whatever way it is defined, will probably always be an ideal that will be impossible to completely achieve in practice.¹⁶⁶ However, the difficulty of drawing the line between public and private reason does not mean that an attempt should not be made. The difficulty

¹⁶² See, e.g., BRANDT, *supra* note 78; HILARY PUTNAM, *REALISM WITH A HUMAN FACE* (1990).

¹⁶³ GREENAWALT, *supra* note 7, at 215–43.

¹⁶⁴ *Id.* at 3–4.

¹⁶⁵ Solum, *supra* note 14.

¹⁶⁶ GREENAWALT, *supra* note 7, at 144–69.

of drawing this line exists not just for the new overlapping consensus, but for any governmental scheme.

Finally, one could object that this Comment has slighted the difference between metaethics and metaphysics. Many of the special problems of metaethics have not been addressed in depth because many of these problems have more to do with the debate between realism and antirealism than with the question of how similar metaethics is to metaphysics. By defining an overlapping consensus that is neutral between realism and antirealism, this Comment has sought to avoid the problems of the action-driving nature of morality and the is/ought gap, which are problems peculiar to metaethics.¹⁶⁷ One of the points of looking at metaethical theories in detail in Parts II and III was to show that many of the problems that separate metaethics from metaphysics are problems of realism versus antirealism, given that many scholars believe metaphysics is objectively real and metaethics is subjective. Part IV demonstrated that government can be neutral on the questions that divide realism and antirealism while still having a principled basis for the exclusion of illiberal ideas.

V. CONCLUSION

In this Comment one simple point has been made, namely that government should not act on ideas derivable solely from intuition. This obtains even if these intuitively derived ideas are part of common sense, tradition, or are approved by a majority. The argument has been conceptualized as an exclusion of intuitionism from the overlapping consensus and a definition of the criteria for public reason as derivable from a coherentist or constructivist theory with a wide circle of justification.

This is not to deny the beauty and the importance of many intuitionist ideas, nor is it to say that intuitionist ideas are false. Intuitionist ideas can be the wellsprings of our individual lives. But intuitionist ideas are uniquely personal in nature because of their very small circle of justification and thus are not suited for a public institution such as government.

Not all intuitive ideas are unreasonable. However, the key to public reason is not reasonableness but public accessibility, and intuitive ideas, because of their purely personal derivation and scope, are not publicly accessible.

Defining the terms of an overlapping consensus does not solve the problem of what actions government should take. It merely defines the arena in which normative theories should be tested and either adopted or discarded. It is only the beginning of the task of defining a system of social justice. This Comment

¹⁶⁷ See *supra* Part II.B.

proposes a definition of the outer limits of the debate. In doing so, it also offers a justified way to exclude illiberal ideas from the overlapping consensus. Thus, this Comment offers a principled solution to the liberal paradox.

This Comment did not attempt to solve the debate between moral realism and moral constructivism. However, in demonstrating that a government can be neutral between moral realism and moral constructivism, the new overlapping consensus proposed in this paper relegates that controversy to the private sphere.

*Susan K. Houser**

* The author wishes to thank Edward B. Foley for his advice in the preparation of this Comment.

